Multi-Sensor Interactive Systems for Embodied Learning Games

Nikolaos Poulios - Anton Eliens

Computer Science. Multimedia Group – Vrije Universiteit Amsterdam

Keywords: multi-sensor systems, educational games, embodied learning, physical interaction, motion interaction, interactive spaces, affective interaction, bio-feedback mechanisms, emotion recognition, interactive theater games, mixed reality

Abstract: This paper explores the use of modern sensor technologies for physical interaction on educational games and interactive spaces. More specifically the thesis studies the potential effect of motion capture and wearable body sensors on educational interactive games, on two aspects: i) on the involvement of human body and motion in the process of learning, and recall of knowledge (embodied learning), ii) on assisting the development of basic social emotional competencies, through the enhanced social affordances of embodied games. The paper presents a prototype of an educational game developed using a motion capture controller and two bio-feedback sensors, proposing a generic architecture for multi-sensor interactive spaces. Finally the paper provides a review of previously studied modalities for emotion recognition, and examines their application on game mechanics and AI game agents.

1. Introduction

The use of computer games in education has been an active field of academic research for the last twenty years, providing considerable evidence to support the positive effects of the use of games on the learning outcomes of students, by increasing their motivation, stimulating their engagement, and by helping students to understand complex concepts by applying them on problem-solving tasks, in an explorative environment. Based on the capabilities of the given technology, up until recently, academic studies focused mainly on the conceptual engagement of learners in games. At the same time, scientists have highlighted the importance of psychological factors that influence children’s learning, placing the development of social-emotional competences at the core of modern pedagogy, and creating a growing need in education for tools and instruments to support and assess these skills.

In the consumer and entertainment industries, recent innovations on sensor and software technology have lead to the development of a constantly growing range of products focused on combining entertainment with physical exercise. Console games designed to be played with peripherals like the Nintendo WiiMote, Microsoft Kinect and Sony Playstation Move/Eye, featuring motion sensors and computer vision algorithms, invite the player to interact in a physical and more active way, using body motion, instead of a conventional game controller.

Other motion and bio-feedback sensors, that were until recently mostly used in medical, sports and training laboratories, have made their way into general consumers, in the form of wearable devices, collecting data during sports or daily activities, allowing the user to monitor and analyze her physical state. Driven by the popularity of electronic games and their ability to engage, a lot of these products use game dynamics, or what is nowadays called “gamification” techniques, like awarding points according to the user’s performance, and giving her the ability to share data and compete with other users-players through
social networks. These techniques aim to engage users more in physical exercise, and to promote personal wellbeing in an entertaining way.

The use of motion controllers and biofeedback sensor technologies has great potential on educational games, contributing to the conceptual engagement of learners, as well as the development of social-emotional skills. Central to this development is the individual in relation to his or her social environment using all possible expressive forms. Motion controllers give us the ability to involve the human body in this process. The human body can be seen as part of the human cognition, and a medium of self-expression and interaction with the environment and other people, thus its participation on learning is of key importance. Additionally, the use of body sensors provide us a way to notice and quantify in real time, the reflection of our actions, or stimuli of the surrounding environment, to our body, physical and mental condition, helping to understand ourselves better, and in extend, to also understand others.

This paper presents a study in the design of learning games for interactive installations featuring multi-sensor technologies, and a prototype developed in order to test and demonstrate the use of such technologies. This study was conducted during an internship at the Waag Society institute for arts, science and technology in Amsterdam, as preliminary research on a project called the Embodied Playful Learning Theater (EPLT). Combining advanced computer vision, motion and wearable body sensors technology, with real time computer graphics projections, active sound and lighting systems, the EPLT is meant to be a highly immersive environment, providing an open platform for research and development of applications and games, targeting to offer a playful and learning experience that will help children to develop their personal, social-emotional skills. EPLT is part of the institute’s involvement in the COMMIT P4, “virtual worlds for well-being” project. The COMMIT+ program is a large, nationwide program, bringing together leading research institutes and hi-tech companies. This research was supervised by the co-author of this paper, Prof. Dr. Anton Eliens, of the Multimedia Group of VU University Amsterdam, and Keimpe de Heer, director of the Creative Learning Lab of the Waag Society institute.

2. Interactive Games in education

Educators continuously face the problem of motivating and engaging their students to learn. Two of the major reasons of this problem are believed to be the passive form of tuition in class, and the gap that exists between learning a theory and understanding its practical value. At the same time younger and younger children are becoming immersed in the consumption of media and the early adoption of technology in their homes.

Following this tendency and observing the high levels of engagement, children demonstrate during gaming, a large number of academic studies during recent years have focused on the design of interactive games for educational purposes. Piaget, through his child development theory believes in the development of cognitive structures through action and spontaneous play [1]. According to Piaget, constructivist learning is rooted in experimentation, discovery and play among other factors. Games provide an alternative, more active and experiential method of learning that supplementary with traditional textbooks can help students understand better complex concepts and engage with content within contexts of use. Balasubramanian and Wilson (2006) analyzed the findings of numerous studies and found that well-designed educational digital games and simulations can help students to obtain critical problem solving and decision making skills, which are necessary for everyday living [2].

Apart from the form of tuition in class, a vital role in its effectiveness is the socio-psychological state of learners. Elias [3], a leading child psychologist, explains the dangers of omitting social-emotional development from classrooms, stating that many of the problems in schools are the result of social and emotional malfunction and debilitation from which too many children suffer, significantly limiting their

http://waag.org/nl/project/commit
http://www.commitnl.nl/
ability to learn effectively. Social and Emotional Learning can be defined as the process of acquiring the skills to recognize and manage emotions, develop caring and concern for others, establish positive relationships, make responsible decisions, and handle challenging situations effectively. Social and emotional learning is of key importance in the pedagogical role of schools, preparing young children to become active parts of the society. Socially and emotionally balanced children have increased confidence, express and communicate better, form better relationships, take and persist at challenging tasks, and have increased capacity of learning.

Goleman [4] outlines five crucial emotional competencies basic to social and emotional learning:

1. Self and other awareness: understanding and identifying feelings; knowing when one’s feelings shift; understanding the difference between thinking, feeling and acting; and understanding that one’s actions have consequences in terms of others' feelings.

2. Mood management: handling and managing difficult feelings; controlling impulses; and handling anger constructively.

3. Self-motivation: being able to set goals and persevere towards them with optimism and hope, even in the face of setbacks.

4. Empathy: being able to put yourself "in someone else's shoes" both cognitively and affectively; being able to take someone's perspective; being able to show that you care.

5. Management of relationships: making friends, handling friendships; resolving conflicts; cooperating; collaborative learning and other social skills.

Despite the image of social isolation electronic gaming has for many people, and the concerns and criticism raised against them by teachers, parents, researchers and policymakers, the literature does not provide convincing evidence to this effect. On the contrary there is a number of studies demonstrating that games often elicit beneficial not only on cognitive skills, but also in affective and social terms (Gunter [5]). Several studies highlight electronic games opportunities for social interaction (e.g. Lazzaro, 2007 [6]) for settings ranging from public interaction (arcades), to semi-public (LAN events), to private (living room at home). In these it has been found that people enjoy playing together or watching others play, sharing comments and enjoying the spectacle and the enhancement of emotional experience that comes from a crowd. When people are playing together, their need to belong is nourished in multiple ways. First, through involvement in a common activity they interact socially, and both the number and quality of social interactions contribute to a person’s sense of belonging resulting in a positive affective state.

De Kort’s and Ijsselsteijn's [7] study denotes the social context effects on player’s performance, caused by the presence of others. The emotional effects include increased arousal, evaluation apprehension, increased self-awareness, self-evaluation, and increased goal relevance. The effects on performance are moderated by the ‘sociality characteristics’ [8] of the game setting, by the other person’s role (co-actor vs. spectator), relationship and expertise, by performance requirements, and by personal differences. Sociality characteristics are the social affordances of the game content, the gaming interface, and the physical environment in which the game is played. Social affordances include the player’s ability to monitor other players’ actions, performance and emotions, and the opportunities for verbal and non-verbal communication.

3. Physical Interaction and the Embodied Learning Theater

As discussed in previous section, children need to be involved in a variety of activities to learn and develop well cognitively, physically, emotionally and socially. These activities include interaction with each other and adults, moving and exploring, manipulating objects, creating representations, listening (and then reading) books, engaged in pretend play, conversing, and building relationships. This
information about children’s needs is the basic reason that early childhood teachers often believe that computers and “screen time” have little place in the early childhood setting; they are correct that technology should not replace these vital experiences of childhood. Rather, technology is most productive in young children’s lives when it enhances children’s engagement in these activities, as well as their reflections about their actions and experiences. The currently prevalent model for educational games in schools is for a single student, or a very small group of students, to work on one computer. This model has limited margins of self-expression and socio-collaborative interactions. The use of modern physical interaction interfaces on hybrid reality interactive spaces for learning will have a great impact on both cognitive and social-emotional engagement of children.

Modern motion capture sensors allow the player to interact with a game using physical movement, map player’s body movement to that of a virtual character, and also to create interactions between virtual and physical objects, by embedding sensors on the last ones. Motion controllers give us the ability to design interactive spaces where physical exercise and social interaction, characteristics of traditional outdoor children games merge with those of modern video games, like rich computer graphics and audio, virtual environments, game dynamics, and interactive story telling.

Murray (Murray 1998)[9] proposes three characteristic values of interactive story experiences: immersion, transformation, and agency. Immersion is the feeling of present in another place and engaged in the action therein. Transformation is the game experience that allows the players to transform themselves into someone else for the duration of the experience. Agency is the satisfying power to take meaningful action and see the results of our decisions and choices. Motion interaction combined with large projections in an interactive space, increases immersion because it gives the player the feeling that he is standing and moving inside the virtual world. Additionally to the feeling of just being present, the player has to follow the action using his body, performing all the necessary moves that the virtual character has to perform, coming to a similar physical state that the character would have in real action, leading to a more experiential, kinesthetic experience with increased feeling of transformation and agency. Previous research comparing motion controllers with conventional interaction controllers, support this claim, finding higher levels of engagement when the controller supports natural movement (Lindley, Couteur & Bianchi-Berthouze 2008)[10]. Another study (Bianchi-Berthouze, Kim & Patel)[11] suggests that body movements appear not only to increase players’ engagement but also to modify the way they get engaged. By inducing body movement, the device resulted in a higher sense of engagement in the players and mediated a feeling of presence in the digital world. The players appeared to quickly enter in the role suggested by the game, and started to perform task related motions that were not required or recorded by the game itself. This supports another factor of engagement, that of fantasy, that often exists on the description of engagement.

Whereas analytical aesthetics is preoccupied with separating humans into mind and body, a part for thinking and a part for sensing, pragmatist aesthetics insists on their interdependencies in the aesthetic experience. In a pragmatist perspective, aesthetic experience is closely linked not only to the analytic mind nor solely to the bodily experience; aesthetic experience speaks to both [12]. Multiple research areas support the embodiment of human cognition, that nearly all cognitive processes are deeply rooted and derived from the body’s interaction with its physical environment (Dourish 2001 [13], Wilson 2002 [14]). Several theorists base this premise on research regarding mirror neurons (Rizzolatti & Craighero, 2004 [15]). Located in the premotor cortex, mirror neurons are activated both when perceiving another’s actions and when producing actions oneself. These neurons are hypothesized to be integral in understanding and imitating the actions of others. The fact that the very same cells are involved in both action and perception suggests that activating potential actions may be an automatic consequence of perception. Based on the theory of embodied cognition, Johnson-Glenberg, Birchfield et al., 2010 [16] suggest the idea of embodied learning according to which learning via movement activates additional
modalities (and sensorimotor systems) for crisper and more stable representations of information. These crisper representations, with more modal associative overlap, will be more easily recalled.

Another important point, highlighted by all studies on motion based controllers, is that controllers that allow natural movement have the potential to offer greater affordances for social interaction. Going back to social-emotional learning, the previous section discussed how the social context effect on player’s performance depends on the social affordances of the game setting, including the game controller.

Aligned with the view of embodied cognition, emotions cannot be seen solely as a mental state but also a physical, bodily, state. Emotions can be generated through imagination without physical interaction, but a lot of previous studies suggest that they can also be generated from physical motion and body posture. A study by Riskind and Gotay [17] revealed how “subjects who had been temporarily placed in a slumped, depressed physical posture later appeared to develop helplessness more readily, as assessed by their lack of persistence in a standard learned helplessness task, than did subjects who had been placed in an expansive, upright posture”.

Long before that, in the field of storytelling, the great theater practitioner Stanislavski, father of socialist and psychological realism, developed a set of training techniques for acting that became established in modern theatre as the Stanislavski system [18]. While in its very earliest stages his ‘system’ focused on creating truthful emotions and embodying them, he later developed the “method of physical actions” to solve the dilemma of spontaneous emotion in a created environment. Main characteristics of this technique during training were improvisation and repetitiveness of physical actions to create a desired emotional response for the character. Emotions were considered to be formed from the subconscious, so this technique allowed the actors to consciously target and control their subconscious emotions through movement.

These previous studies allow us to consider body motion and gestures, designed to interact with a game, as another modality to stimulate player’s emotions and create a stronger experience. During this experience, the player has the freedom to express her emotions and communicate with co-players or spectators using her whole body and motion. This motion gives spectators a theatrical perspective of the game. All the in-game action becomes visible; spectators can monitor player’s performance, physical effort invested, and expression of emotions. The player becomes active part of the story, and a bridge between the physical and imaginary dimensions of storytelling.

Through this improvisational performance, player’s evaluation apprehension is increased, along with self and others awareness. Viola Spolin [19], another innovator of modern theatre, used the element of improvisation, since her early work, developing theater games for children and adults, that unleashed creativity, adapting focused “play” to unlock the individual’s capacity for creative self-expression, through self discovery and personal experiencing.

In addition to motion controllers, this study examines the application of wearable biofeedback sensors, as another interaction technology to enrich game play and assist in social-emotional skills learning. Nowadays there are a lot of commercially available, ready-made sensors, as well as kits and platforms to build custom ones. The in game use of devices like heartbeat, body heat, and brainwave sensors, can provide a quantified image of the player’s physical and mental state. This information can be also seen as an additional tool to interpret player’s emotional state.

The real time visualization of the collected sensor data, and their use inside game dynamics, have great potentials to increase engagement and awareness of both players and spectators. Between players, this information increases their self-awareness, and motivates more active participation inside the game, fostering also competition with others. At the same time spectators are also more engaged into the game, given an augmented view of the player’s performance, her physical effort and responses to the game’s environment stimuli.
4. A prototype for a multi-sensor interaction physical learning game

As preliminary research for the EPLT project of the Waag Society institute, this study focused also on the technical aspects of designing a multi-sensor system for game installations. After a series of smaller prototypes testing capabilities of some single commercial sensors, the aim was to develop a prototype game featuring multiple sensors interaction. Targets of this prototype were: i) the testing and demonstration of the capabilities of some particular commercial sensors, ii) the basic implementation of a multi-sensor system's, generic architectural design, and iii) to provide a base for the collection of some first body and motion data for further study of the concept of using these data in game interactions.

Main idea of the prototype conceptualized and developed by the author was to use a motion capture sensor to create a board game that would blend traditional forms of children games with modern video games. Characteristics such as dynamic computer graphics, sound effects, fantastic virtual worlds, and the ability to play with someone over distance have made video games very exciting and engaging to children. On the other hand, traditional games that used to be more popular in the past, for example hopscotch or jumping rope, although they might seem a bit unsophisticated for today's hi-tech standards, motivated the physical exercise of children while offering a playful experience and an opportunity of social interaction. Modern devices such as the Microsoft Kinect sensor, give us the ability to combine the best parts of both forms of games.

4.1. The technology

The final prototype presented here uses the Microsoft Kinect sensor, the Neurosky Mindwave EEG sensor, and the Zephyr HxM ECG sensor. As this research did not focus on the development of custom sensors, the selection of these was based on their availability as ready-made solutions, their technical specifications, the level of support in the programming development of the game, and their suitability to be used in a physical installation.

**Microsoft Kinect** is a real-time motion capture sensor developed for the Xbox game console. Based on an infrared depth camera, and advanced motion machine learning, computer vision algorithms [51], the Kinect sensor is capable of tracking at 30Hz the position of 20 joints of the body in space, for two players simultaneously. Kinect was chosen as a state of the art commercial motion sensor, an economic solution with previously proven performance and reliability.

**Neurosky Mindwave** is a single electrode wireless, EEG based sensor, designed mainly to be used in games. Mindwave places an sensor to the forehead of the player, capturing voltage fluctuations in specific frequencies that have been related to brain activity, providing as output two values of “attention” and “meditation”, that indicate the mental state of the player. An EEG sensor was chosen as a new in commercial level technology, with promising features that would add novelty to the game, and will intrigue the player. Mindwave was selected specifically for its easiness of use in an installation, requiring short time to wear and calibrate, and the ability to maintain position and signal quality under intense motion.

**Zephyr HxM** is wireless device that places an ECG sensor near the chest of the player, using an elastic band. The sensor provides the heartbeat rate of the player, and it also includes an accelerometer based on which the sensor can calculate the pace of the player. Heartbeat rate was chosen as a value that in contrast with the EEG brain activity signals, all people and even children are quite familiar with, because large enough fluctuations are internally sensed by our senses. The visualization of the heartbeat rate, and the interaction based on it, creates another link between the physically sensed body, and the virtual environment, that will enhance the feeling of immersion.

The combination of these two selected body sensors creates a good first base to collect data and study possible indications of correlation of physical activity and mental state.
The prototype was developed using the Unity 3D game engine for the game, OpenNI SDK for communication with the Kinect Sensor, and cinder C++ library for communication with the other sensors, visualization and storage of the data collected. All game artwork was taken from Unity's community and example projects.

4.2. The Game – NumHop

In the final game prototype developed, named NumHop, the player is placed on a virtual large hall [Figure 1]. In front of the player placed on the floor, there is a board of 16 numbered tiles. The player is called to answer questions on simple multiplication matrices, for example the result of 6 x 7. The tiles of the board are numbered to values close to the correct result with at least one containing the correct number. The player then has some seconds to select his answer by stepping onto a tile. The faster the player responds correctly, the more points he gains. If the player does not respond, faster the player responds correctly, the more points he gains. If the player does not respond, the game moves automatically to the next question. If on the other hand the player selects a wrong answer, the board moves to the next question, and an enemy robot is teleported in the scene through one of the 6 chambers, and starts approaching the player with bad intentions. The player can defend himself against the robots by activating his “superpower beam” (activated by raising both hands above shoulder level) and aim a target against the robots. The player starts the game with a certain level of superpower that it is reduced by use. When however the Mindwave sensor that player wears on his head, detects high level of attention, the superpower level starts to increase and the player can use it again. If the player runs out of superpower, he has to suffer the robot’s hits, which reduce the player’s health level. If the player survives an attack he can step back for a moment and try to relax. When the Mindwave sensor detects high level of meditation, the health level of the player is increased. The player is given 3 lives in the beginning of the game, and bonus lives be awarded after a number of consecutive correct answers.

The GUI of the game includes four circular meters. The first one (purple) indicates the current meditation level, the second (green) the level of attention, the third (cyan) indicates the level of remaining superpower, and the third (red) indicates the in game characters health level, and the real heartbeat rate of the player in beats per minute. The heart rate value is not directly connected to any element of the game. Although there was the idea of correlating the heart rate to the update interval of the board, it was finally abandoned. There are two reasons for that. The first one is that because the heart sensor has to be placed under player’s clothes, it might be proved impractical to use on a school test environment. The second one is that designing a certain interaction based on heart rate requires to know in advance the expected range of values during the game, knowledge and expertise that was not available at the time of development. The presence of the hear rate value however was thought to be useful as also explained earlier, first for observation of the values for further use, and second to see how players respond to this information, if for example by placing the heart rate value of the player, appearing on the GUI, as what could be conceived by someone as another form of score points, motivates the player to raise his heart rate by moving more intensively.

4.3. Prototype Evaluation

The prototype was tested in a series of private sessions and an open evaluation session that took place at the Theatrum Anatomicum, at the Waag Society, with participants from people working for the institute, and other volunteers. Overall 12 people tested the prototype, from which 4 people played the game using only the Kinect sensor, 5 people using the Kinect and Mindwave sensor, and three people using all 3 sensors. The evaluation data where collected either by a questionnaire or by short interviews and discussion with the participants after they played the game.

Overall the prototype presented was rated as a fun and very original gaming experience. The prototype’s concept of motion-based virtual board games was rated to have high potentials for learning games. The motion interaction with the Kinect sensor was rated very exciting, and specially people who did not have played another Kinect game before, were amazed by the capabilities and potentials of the device, and although the game avatar’s
movement was lacking the physical human kinetic motor models found in commercial game titles, this did not seem to frustrate players, since the responsiveness and robustness of the sensor were very high. During the test the players where not explained in detail how the superpower beam works. The only tip given was that it can be activated by raising both arms above the level of the shoulders, and that it stays activated as long as the arms stay in that level. The activation method of raising both arms appeared to work good for the game, being easy to understand both for the player and the sensor, and not very strict to limit players natural movement. In most cases players quickly found out how to direct the beam by rotating their body, and found this action enjoyable to perform.

The Mindwave sensor also proved to work good as a piece of hardware for the game. It was rated as quite comfortable to wear, it was easy enough to get the quality signal, required by the sensor for the attention and meditation level determination to work, and maintained it without problems through the game. In some cases the cases would lose signal after a jump, but it regained it shortly, without the player stop moving. The metaphor of connecting attention level to the superpower level, and the meditation to health level was rated very high as a concept; the effectiveness of the sensor inside the game though, did not receive very high ratings. The attention and meditation values usually stayed on an average neutral level through the game, and in a lot of cases we had to cheat by adding superpower level by the keyboard, in order not to spoil the fun. If a player dedicates more time to master the sensor, results will get better, but spontaneously during the game, the calculations did not seem to raise the attention levels that much for the sensor to play its role on the game very effectively. It should also be noted though, that all participants in the evaluation were adults, and basic multiplication matrices are not so challenging for them. This leaves an open possibility that for young children, for who the prototype is designed for in the first place; these calculations might be a heavier mental load, thus more easily detected by the sensor, raising the effect of the sensor in game play. In any case the use of the sensor certainly adds some excitement and curiosity about the game, since most people are not familiar yet with brain computer interface devices, and they want to try and learn how they work. All participants that played the game without using the mindwave sensor, were very curious to try it with the additional sensor, and believed that this would certainly add more fun to the game.

As mentioned earlier, the heart beat rate was not given an actual role on the game, because it was believed that the use of the heart sensor might be problematic during an evaluation session, basically because the player has to wear it under his shirt. Indeed given the option, most participants chose not to use the heart sensor (5 out of 8). Nevertheless, the ones who used the sensor rated it as very comfortable to
wear, and the sensor worked almost perfectly through the game. Since that session was made to test the game and the sensors, participants who used the heart sensor were monitoring their heart beat rate, and felt motivated to try and raise it, in order also to test the responsiveness of the sensor and their condition. Again the integration of the heart sensor on the game is a first step to gather and study heart rate values and ranges, in order to get some knowledge based on which other interactions and game dynamics can be developed. Additionally new trends in the use of body sensors in daily life push the technology around sensors and we already have samples of sensors embedded to ordinary accessories, like for example a wrist watch with a heartbeat sensor, and techniques to measure heart rate from a close distance using infrared light. Devices like these will make the integration of biofeedback sensors to games easier and more practical to use.

The prototype, as such, followed a short development cycle. All graphics were taken for freely available resources on the Unity game developing community. This leaves a lot of room for changes and improvements to make the game environment more suitable and pleasant for younger children. Introducing additional motion based interactions could expand the concept of virtual worlds in a board game. For example, the board could be designed additionally as a floating platform, kind of a flying carpet, which the player can navigate through space depending on his position on it, called to follow a track path and encounter enemies on his way to finish the level. On a more permanent setup of a game installation, the feeling of immersion inside the games world could be enhanced by introducing additional interactions with vision and sound, for example in the prototype presented, the room could be lighted with intense red light whenever the player gives a wrong question, and a robot attack is launched, or intense blue lights whenever the player’s lighting superpower is activated. In a similar way, the game could change the color of the lights when necessary, to help the player relax for a while, in order to charge his health level and reduce his heart beat level.

The game developed is using one motion capture sensor and two body sensors, one of which did not have a specific role in the gameplay. After a short presentation about the sensors and the game, and before playing it, some people did not understand exactly how the body sensor values are used inside the game. This fact could mean that if the game was using more sensors, with more complex game dynamics, and if it was presented to young children, the game interactions could be hard to understand. On the other hand, all computer games require from the player to invest some time playing in order to discover all the game’s mechanisms. As wearable sensor technology develops, making them more practical to use, the greatest challenge for a multi-sensor game designer is to create a meaningful ambient interaction layer, through which the player will discover and experience the game’s mechanisms while playing, rather than require a detailed explanation in advance.

5. Multi-sensor interactive system architecture

NumHop prototype was developed following a design pattern that demonstrates a generic architecture for multi-sensor interactive spaces, such as the EPLT project. As described in the introduction of this document the EPLT is meant to be an open platform to be used by developers, artists and researchers for the development, experimentation, testing and support of multi-sensor technologies applied on interactive applications. As such, the EPLT should feature a flexible, extendable and scalable architecture that can be adapted according to the application built upon it, and the equipment used for input and output of the interactions.

Based on the above system requirements, characteristic of the proposed architecture is the separation of the system and its interactions in three levels (Figure 2), the first level is the world level including the physical setup of the installation, the sensor devices used, and the various output devices of the system, such as projectors, sound and lighting systems.

The second level is the device level, describing low-level hardware and middleware responsible for the collection and transmission of data from the various sensors to the application, and the sub-systems controlling the output mechanisms used by the application. The last level, the application level corresponds to the system accepting data from the sensors as input and process them to the corresponding output. Separating the application from the device level, allows the development of a common middleware that can be easily
extended and used in multiple applications. Components composing the different parts of the device and the application level can correspond either to processes running on the same computer, or to processes running distributed over a network of computers, each one implementing a different part of the interactive system. In order to provide this scalability, a messaging service is established between the device and the application level.

Following this architecture, NumHop consists of two applications, the game itself developed in Unity game engine, and a second one, developed in cinder C++ library, implementing the device level. That second application is responsible for the connection with sensor devices through Serial over Bluetooth, the collection, over time visualization, and permanent storage of the sensor data, along with the transmission of the data to the game engine. The two applications communicate using the Open Sound Control (OSC) protocol. OSC is a communication protocol originally developed at the UC Berkeley Center for New Music and Audio Technology (CNMAT), for communication between computers, sound synthesizers, and other multimedia devices, optimized for modern networks.

OSC's advantages include interoperability, accuracy, flexibility, and enhanced organization, featuring open-ended, dynamic, URL-style symbolic naming, symbolic and high-resolution numeric argument data, pattern matching language to specify multiple recipients of a single message, high resolution time tags, and “bundles” of messages whose effects must occur simultaneously. OSC messages are usually transmitted over the UDP protocol. Due to its flexibility and simplicity, OSC has become an “industry standard” in the field of interactive installations, and has been implemented in a lot of programming languages and libraries, a growing number of real time sound and media processing environments, software and hardware synthesizers, sound and light consoles and various tangible interfaces. This popularity of OSC makes it an ideal solution to be used in the system’s middleware, as it allows the interaction and experimentation with sensors, even by non-programmers. Messaging between layers allows sensor values to be transmitted to multiple systems at once, keeping them independent from each other during implementation and while functioning. As an example, apart from the output produced by the game, sensor values can be directly mapped
through OSC to channels of a sound console or a DMX lighting controller, triggering additional sound and light parallel interactions. This makes the system easy to extend and allows someone, or a group of people, to design interactions using a variety of software and hardware that as a whole will create a rich and immersive interaction experience.

Another advantage of having a different process controlling the communication with the sensors is that allows us to develop an input control system that makes it easier for the system to handle and recover from errors. During a game session, a sensor might suddenly lose connection with the system or lose contact with the body of the player because of a sudden move or a jump, resulting in incorrect or no values. Errors like that are easier to spot by someone monitoring the signal qualities and raw values of the sensors. In case of an error, the staff member can temporarily disable the transmission of sensor data to the rest of the system, leaving it to continue running based on a previous valid state, while trying to reset the connection with a particular sensor, or if the error is critical for the progress of the game, he can decide to pause the game, help the player to get a sensor placed correctly on his body, and resume the game.

6. Biofeedback mechanisms and Affective Interaction

In the previous sections we described how emotions can been seen as a combination of mental and bodily state, we saw the example of how a simple EEG sensor can provide us with a basic indication of arousal and relaxation levels, and through the NumHop prototype, how sensor information can be used to design simple interactions that will increase engagement and social affordances of a game. Diving more into the concept of games for developing of social-emotional competences, and towards more complex and meaningful bio-feedback interactions, this section aims to provide a brief review of all previously studied modalities related to emotion expressions, and study their application for the development of artificial emotion intelligent agents in games.

The vision of machines with emotional intelligence coexists with that of artificial intelligence since the invention of the term. It is a popular theme in science fiction literature, featuring androids understanding emotions and having human like behavior, and aptly raising ethical questions about the use of such technologies. Picard (1995) [20] first coined the term "Affective Computing", describing interactive systems that have the ability to interpret the emotional state of users and adapt their behaviour to them, simulating human empathy. Although we are still quite far from this vision (or nightmare for some), research laboratories around the world work on developing emotion-sensing technology to support the study of human behavior, the affective human computer interaction, and communication between people, with a broad range of applications including psychology research, computer assisted therapeutic systems, safety monitoring applications, assessment and training systems, user experience studies, marketing research, and automatic affect-based indexing of digital material.

As we know from personal experience, emotions are hard to define and recognize. Despite all our senses, the verbal and non-verbal communications skills we have as humans, it is often hard to immediately recognize someone’s emotions. Expression of emotions is becoming even more complex when analyzed in a global, cross-cultural scale. It is easy to imagine thus, that emotion recognition is a very difficult task for a computer, especially on real time application where the system has to analyze the user’s state and give a response on a very narrow time frame. Classic psychological research claims the existence of six basic expressions of emotion that are universally displayed and recognized: happiness, anger, sadness, surprise, disgust, and fear, other studies on emotion recognition also include emotions like despair, interest, irritation and pride [21]. A lot of studies do not accept this categorization of emotions, suggesting that it is not emotions but some components of emotions that are universally linked with certain communicative displays. Most theorists agree that the two dominant dimensions of emotion can be described as valence (pleasant vs. unpleasant) and arousal (activated vs. deactivated or excited vs. calm) [22]. Mapping even basic emotions
on these two dimensions is challenging [Figure 3.], and emotion recognition systems analyzing a single human modality usually suffer either from poor accuracy or over simplified classification of emotions. For this reason a lot of researchers avoid the term “emotion recognition”, preferring that of “biofeedback mechanisms”, criticizing what Boehner et al. [23] call “Informational approach to emotions”, in which emotions are represented as clear, discrete states, in a machine-readable format. Boehner suggests an alternative view called “Interactional approach”, where instead of trying to develop systems to recognize human emotions, it focuses on helping humans understand, experience, and express their emotions through technology.

Nevertheless, considering both approaches, and as also described in previous sections (Sect. 3), the application of multi-sensor interaction in a gaming installation will contribute to the gaming experience in various ways. Research has shown that video games can stimulate strong emotional reactions from players, making the an appropriate field for further research on the relation between bio signals and emotions, on affective interaction and behaviour studies.

The application of multi-sensor biofeedback mechanisms in games can assist in learning mood management techniques, similar to techniques used in professional and sports training. Biathlon for example, which combines cross-country skiing with rifle shooting, requires special techniques from the athletes to control their breathing when they arrive to the shooting range, after a very demanding physical effort, and with a very high heart beat rate. Crews and Landers (1993)[24] identified electroencephalographic signal (EEG) measures of intentional patterns prior to successful golf putts. Pope and Stephens (2011)[25] describe how the concept of physiological modulation of operator input, evolved from a physiologically-adaptive simulator system that was developed in National Aeronautics and Space Administration (NASA) flight deck research. In this system, EEG signals of pilots controlled the level of automation in a simulator flight deck. This “closed-loop” testing setup was used to determine what level of automation kept pilots best engaged in the flight task. It was soon realized that, given enough practice, pilots could probably turn the testing system into a training system; that is, they would learn to control their EEG to set the level of automation where they preferred. This becomes essentially an EEG biofeedback training situation. In a similar way that games based on motion sensing controllers reward players for imitating a skilled performer's overt motor behavior, biofeedback

Figure 3 Emotions mapped on basic dimensions
mechanisms can additionally challenge the player to reproduce the expert performer’s emotional and cognitive state by setting as a target the psycho-physiological responses exhibited by the expert in a real-world situation.

Bio-feedback mechanisms in games have also great potential in developing virtual actors demonstrating basic artificial emotional intelligence. These virtual actors can for example motivate players and reward physical effort, or help them to calm down. Intelligent virtual actors can utilize emotion recognition mechanisms in certain points of the story, asking players to act emotions or behaviours, or in order to perceive players’ reactions to game stimuli and trigger virtual actor’s corresponding behaviours, e.g. simulating empathy. Elements like these would increase the engagement and enhance the experience of an interactive story where the player finds herself in an immersive world, inhabited by personality-rich, robustly interactive characters.

The next part is a presentation of the various modalities used to capture physiological signals that can be associated with the emotional state of a person, along with software for emotion recognition developed from previous research.

6.1 External modalities

Speech analysis

Speech is the primary method of human communication. Analysis of certain features extracted by speech characteristics like intensity, pitch, phonetic features, voice segments, pause length, and spectral modeling, along with linguistic analysis based on keywords used, can be used to make conclusions over the emotional state of a person [26].

Facial expressions

Facial expressions analysis has been the first, and extensively used since then, method for emotion recognition on multiple studies, and it is the preferred method for single modal emotion recognition systems. Facial expressions are the main non-verbal communication tools, providing the most powerful, versatile and natural means of communicating motivational and affective state. Apart from expressing emotion, facial expressions are providing important communicative cues during social interaction, such as our level of interest, our desire to take a speaking turn and continuous feedback signaling understanding of the information conveyed. Facial expression constitutes 55 percent of the effect of a communicated message [27] and is hence a major modality in human communication. Several studies have also shown that ordinary people can detect six emotional facial expressions with an accuracy ranging from 70% to 98%.

In facial expressions analysis systems, the face is segmented focusing on the facial areas of eyes, eyebrows, mouth and nose. Each of these feature-candidate areas contains the features whose boundaries are extracted and stored over time, and then the displacement of each feature is compared to “neutral face” model images to conclude the emotion expressed by the subject.

Body movement/postures

Although a lot has been written for the so-called “body language”, body movement and posture has not been researched on emotion recognition so extensively, as facial expressions and voice analysis. There are though some studies, questioning the validity of facial expressions as a modality for recognizing affective states, because face is involved in various functions and many of the famously recognized facial expressions represent only a small subset of the possible expressions, suggesting body posture as a very good indicator for certain categories of basic emotions. Most studies however, have not been able to demonstrate similar recognition accuracy with that of facial expressions classifiers, especially those who study emotion recognition from static body postures only. Coulson [28] considered how 6 joint rotations (head bend, chest bend, abdomen twist, shoulder forward/backward, shoulder swing, and elbow bend) could help recognizing 6 emotions (angry, fear, happy, sad, surprised and disgust). Concordance rates for attributions of the 6 emotions ranged from zero for many disgust postures to over 90 percent for some anger and sadness postures. Kleinsmith A. and Bianchi-
Berthouze [28] used four affective dimensions (valence, arousal, potency, and avoidance) instead of discrete emotion categories. On their study there was a 12% error percentage for valence, 10% for both arousal and potency, and 11% in the case of avoidance. In their conclusions they report that other types of body motion features, may be necessary for achieving better recognition of some affective states such as fear, and better performance of their model. Other studies that include body motion as a modality [29], tracking features like quantity of motion and contraction index of the body, velocity, acceleration and fluidity of the hand’s barycenter, orientation and approach/avoidance behaviors of two participants towards their interlocutor in an interaction, suggest that body language reflect their level of activation and dominance but are less informative about their valence (positive vs negative).

Pupil size

Studies have shown that the eye’s pupil is significantly larger during both emotionally negative and positive stimuli than during neutral stimuli [30]. Although we cannot distinguish valence, pupil size can be used as an additional modality of arousal. A lot of eye tracker devices have the ability to measure the pupil’s size.

6.2 Internal Modalities

Emotion recognition systems based on external modalities like the ones described above are more familiar to use, because they accept the same input we as humans do, in our everyday interactions with others. The performance of such systems however, depends on environmental conditions, and the training models that have been used on their machine learning algorithms. Although advanced processing algorithms have been developed to minimize the effect of environmental conditions, training classifiers can be practically difficult and very time-consuming procedure. In other cases, and special groups where there are difficulties to communicate expressively, the above modalities are also not effective for emotion recognition. For example, on the autistic spectrum, where an autistic person might outwardly appear calm and relaxed, while experiencing a state of emotional or cognitive overload. For these reasons scientists have also turned to the use of embodied biophysical sensors, monitoring signals that can reveal valuable information, not only for the physical state of someone, but for the emotional and mental state as well.

The physiological signals usually monitored in behavior studies are:

**Heartbeat rate (ECG)**

Electrocardiography sensors determine heartbeat rate by detecting and amplifying the tiny electrical changes on the skin that are caused when the heart muscle depolarizes, by measuring the difference in voltage between two electrodes placed either side of the heart. There are also optical heartbeat sensors using an infrared LED and a phototransistor, placed closed to each other with usually a fingertip, or the ear lobe, in between. These sensors work based on the fact that when your heart beats you have a quick rush of blood into tiny blood vessels close to your skin, which makes it less transparent, so less light comes through it to the phototransistor. Changes in heartbeat can give us a clear index of arousal, but sensors are prone to movement artifacts. Increase of hear rate has been related to fear, and decrease to anger [31]

**Galvanic Skin Response (GSR)/Electro Dermal Activity (EDA)**

Both terms refer to the electrical changes measured at the surface of the skin. EDA sensors usually work by passing a miniscule amount of direct current between two electrodes in contact with the skin. When a person experiences emotional arousal, increased cognitive workload or physical exertion, the brain sends signals to the skin to increase the level of sweating. Sweat is a weak electrolyte and good conductor, the filling of sweat ducts results in increasing the conductance of the applied current. Changes in skin conductance at the surface thus provide a sensitive and convenient measure of assessing sympathetic arousal changes associated with emotion, cognition and attention.

**Skin temperature/Heat flux**

Refers to the amount of heat that the body emits. Studies have shown that Heat
Flux is effective in detecting context switches. This is because context switches often involve physical movement, which causes the body to warm up and therefore emit heat. Heat flux has been reported to increase during increased cognitive load.

**Electroencephalography (EEG)**

EEG is the recording of electrical activity, using electrodes attached along the scalp, measuring voltage fluctuations resulting from ionic current flows within neurons, and generated by the synchronous activity of thousands or millions of neurons with similar spatial orientation in the brain.

EEG has been widely used in clinical research, on neurology, to diagnose epilepsy, coma, brain death and various encephalopathies. Scalp EEG activity shows oscillations at a variety of frequencies and researchers have associated certain oscillations frequency ranges and spatial distributions to different states of brain functioning. Usually three frequency ranges are used for this purpose: i) Theta (4 - 7 Hz): related to drowsiness, ii) Alpha (8 - 13Hz) related to relaxation, iii) Beta (>13 - 30Hz) related to alertness.

Although EEG is not the most accurate method to monitor brain activity, its ease of use, portability and low set-up cost has made it the most studied one, and resulted its application to other research fields and all kinds of experiments where it is interesting to monitor the mental state of the subject. During the last years, EEG has made its way to human computer interaction research, and a small number of companies are working on developing low cost, non-invasive, brain computer interface products like the Emotiv headset and Neurosky’s Mindwave. These interfaces are currently used mainly in gaming and other entertainment applications, since they are still proved to be inaccurate and not practical for more critical applications.

**Functional near-infrared spectroscopy (fNIRS)**

fNIRS is another technique for sensing brain activity, similar to the technique used by optical heartbeat sensors mentioned earlier. The fNIRS system is made up of probes that send light at two wavelengths in the near-infrared range. The main absorbers of the light are oxygenated hemoglobin and deoxygenated hemoglobin. These act as relevant markers of hemodynamic and metabolic changes associated with neural activity in the brain. The reflected light is then picked up by the detectors on the device. Depending on the amount of light that is reflected, we can get a measure of brain activity in the area beneath the sensors.

Studies in fNIRS [33] report that the hemodynamic response being measured in brain is a slow response which occur over 5-8 seconds. This makes the technique currently impractical to be used for interaction input interfaces. For the moment there is still no commercial brain computer interface utilizing the fNIRS technique.

### 6.3 Biofeedback Interactions.

**Thoughts and insights**

The biofeedback mechanisms used in a game are defined by the interactions designed for it, on the questions of what we want to measure, why we want to measure it, and how we use this measurement inside the game. However, designing biofeedback interactions for a game installation featuring also physical motion interaction, automatically sets some factors of the game setting that have to be considered.

Bio-signals collected from sensors can be used in game interaction basically in two ways: i) as a continuous monitored signal, correlated to running variables of the game, such as the difficulty or pace; or monitoring player’s progress towards a desired state, perceived as goal; ii) as signals monitored in relatively small time frames, on specific points of a story-driven game, acting as sensing mechanisms of virtual agents. Certain signals like ECG, body temperature/thermal flux, and EEG, are offered more to be handled as continuous. Naturally, in a game with intensive physical motion interaction, values of heart rate, temperature, and skin conductance are expected to increase, which could make their use for emotion recognition purposes problematic.

**EEG signals** are also interesting to monitor through a game. EEG sensors can give an indication of the cognitive load of the player, making it interesting to study the
correlation of physical and mental state during a game, seeking additional signs to support the idea of embodied learning. The main problem with EEG sensors is that the signals monitored are very weak and noisy, and the electrodes must be positioned very carefully on scalp.

**Pupil size** was found not practical to use as a biofeedback mechanism. An immersive environment with continuous visual stimuli and physical motion of the player is expected to affect eye movement and pupil size, and the requirement of a wearable camera in close distance from player’s eye, will limit the sense of freedom to move.

As mentioned earlier, accuracy of emotion recognition systems based on **facial expressions, speech analysis and body postures** can be limited by environmental conditions. Thinking specially of a game with motion interaction we would expect the user to move a lot inside space, being in a distance of some meters from the camera, and adopting postures suggested by the game action. These factors, along with fast transitions from one emotional state to another, experienced during intense moments for the progress of the game, make these modalities practical to monitor as continuous signals only for later statistical analysis, and not as signals correlated continuously with runtime properties of the game play.

**Skin conductance** value has been reported to vary a lot between persons being in relevant states, and detection of sudden emotional context change is noticed as sudden increase of the value compared to previous ones. This makes skin conductance also more suitable to be monitored on a short timeframe where we want to monitor player’s reaction to specific game stimuli.

The ultimate application of biofeedback mechanisms in story driven games, would be to develop virtual actors demonstrating signs of artificial emotional intelligence, by reading input from sensors while interacting with the user. As an example, Self City [34] is a previous project of the Waag Society on gamefying social-emotional skills learning. Self City transferred the player into a virtual city in which the player could train her social skills by interacting with other virtual avatars, in simulations of daily social life, and conflict scenarios. On these scenarios the player was called for example to deal kindly and calmly with an aggressive doorman, or someone who took his place at the tickets queue, or kindly ask another person for something. The player was guided by another avatar, her personal social skills advisor. Self City was designed on Second Life, an online virtual world in which users interact with each other through avatars. Behind all avatars of Self City there where educators interacting with the user.

Transferring game scenarios like those of Self City to a multi sensor interactive space, virtual actor could use emotion recognition systems to sense player’s emotions and trigger corresponding responses from a pre programmed behaviour library. A virtual actor could enable emotion recognition at the beginning of an interaction, for example when the player is in a close range, sensing if the player is for example kind, smiling (facial expressions), talking calmly or using “please” (speech analysis/recognition); or if the player looks angry or if she is scared (skin conductance). Sensing the above modalities the virtual actor could detect the end of a phrase or a pause, and use the output of emotion recognition algorithms running, to trigger behaviours based on the story script, awarding the player, simulating empathy or act like it has been insulted or upset. Although current state of the above technologies may require from the player to overact her emotions, and delayed responses would create an unnatural flow in the interaction, considering the progress made for example on speech and action recognition systems during the last years, implementations of intelligent virtual actors will become more appealing and easier for interactive story telling.

7. Conclusions

This document makes an effort to build a theoretical framework around multi-sensor interactive spaces for educational games. Combining results of previous studies in the fields of educational and commercial games, psychology, theater and kinesiology, provide us with encouraging findings to support the potentials of motion capture and biofeedback technologies in learning games. Inspiration for such games can be find not only in story driven games, but also in more free play, traditional
children games as well as improvisational theater games that have been already used in schools. The elements of body motion and the reflection of action through biofeedback mechanisms will increase conceptual engagement, offer more opportunities for social interaction and assist the development of social emotional competences, and creativity of children.

The experience of the prototype evaluation showed us that these technologies add a certain level of novelty to the game, triggering people’s curiosity and offering and engaging experience. Even if these technologies are becoming more mainstream in households and general consumers, which for now is the case for motion controllers, these still find their essence in the immersive environment of an interactive space. In such an environment the player have more freedom to move and express, and become more engaged to their performance, leading to an affective kinesthetic experience. For spectators who can monitor player’s performance, he becomes an active physical part of the game creating a theatrical atmosphere around his performance and the game's story and virtual environment. As players and spectators exchange roles the overall experience leads to increased self and others awareness.

The experience of NumHop also showed us that even in a simple game, designing interactions based on multiple sensors might lead to difficult to understand game mechanics, specially since most people are not that familiar with these technologies yet. Additionally some of previous studies using sensor data lead to an oversimplified interpretation of an emotional state. One suggested way to overcome this problems is to follow a different, more ambient approach of interaction, where sensor data is not directly mapped to certain game play variables, but contribute to the creation of an atmosphere in order to motivate players self exploration and free expression. The other approach is to use multiple modalities in order to get a better picture of the players state, and use in game material and virtual agents that will interact with the player and help him explore and understand better the game interactions.
References


